

2019/2/26 14:20-15:20

日本薬剤疫学会 薬剤疫学とデータベースタスクフォース  
実務者のためのデータベース研究講座 その1

# 国立病院機構のデータを用いた validation studyの実際

山名隼人

東京大学大学院医学系研究科ヘルスサービスリサーチ講座  
国立病院機構本部

# 本日の内容

- **Validation study**とは
- **国立病院機構(NHO)のデータを用いた validation study**の実際  
*J Epidemiol 2017;27(10):476-82*
- **NHOのデータを用いた臨床疫学研究**の実際

# 診療情報の活用

## 診療情報

- DPCデータ・レセプトデータ
- 電子カルテデータ（検査値等）

## NHOにおける活用法

- 診療機能分析レポート
- 臨床評価指標
- 病院経営
- 臨床疫学研究



# 診療情報の活用

## 診療情報を用いた臨床疫学研究

- 大規模データベース
- 観察研究 ↔ RCT等の介入研究
- 既存のデータを二次利用 前向きにデータ取得
  
- 実際の臨床現場を反映、“real-world”
- 観察研究からも重要なエビデンスを生み出すことができる

# 診療情報の活用

## 臨床疫学研究の例

(DPC調査研究班)

1. 下肢切断患者における術後死亡・再切断のリスク要因  
Endoh S et al. *Progress in Rehabilitation Medicine* 2017; 2: 20170015.
2. 市中肺炎に対する抗生剤のde-escalationの安全性  
Yamana H et al. *J Infect* 2016;73:314-25.
3. 重症熱傷に対する気管切開の効果  
Tsuchiya A et al. *Burns* 2018;44:1954-61.
4. 心肺補助を要するショック患者に対する  
早期経腸栄養の影響  
Ohbe et al. *Intensive Care Med* 2018 44:1258-65.

# 診療情報の質

## データの妥当性 (validity)

- 査読者コメントで多いもの：“Is the data validated?”
- データベースに格納されている情報と、研究者がイメージする情報の差
- 100%の正確性は現実的ではない
- 定量化することで、活用の質を向上させる
- ~~診療、記録・レセプトのチェック~~
- ~~正確性の評価・ペナルティー~~
- ~~活用を目的として日々のpracticeを変える~~

# 妥当性の指標

## 感度・特異度

至適基準（gold standard）との比較が必要

例：「脳梗塞」の病名

- 至適基準 = カルテ上の病名
- 検証したいもの = データベース上の病名

の場合、

- 感度：カルテ上、脳梗塞がある人のうち、データベース上の病名でも脳梗塞ありとされる人の割合
- 特異度：カルテ上、脳梗塞がない人のうち、データベース上の病名でも脳梗塞なしとされる人の割合

# 妥当性の指標

## 感度・特異度

		カルテ上の脳梗塞		
		あり	なし	
データベース上の脳梗塞	あり	真の陽性 a	偽陽性 b	陽性的中率 $=a/(a+b)$
	なし	偽陰性 c	真の陰性 d	陰性的中率 $=d/(c+d)$
		感度 $=a/(a+c)$	特異度 $=d/(b+d)$	

- 感度90%, 特異度90%, 実際の割合5% → 陽性的中率は32%



# 妥当性の指標

## 感度・特異度

		カルテ上の脳梗塞	
		あり	なし
データベース上の脳梗塞	あり	真の陽性	偽陽性
	なし	偽陰性	真の陰性

- 感度90%, 特異度90%, 実際の割合5% → 陽性的中率は32%

# Validation studyの必要性

## 感度・特異度の研究結果への影響

- A薬 vs B薬でアウトカムを比較したい
- 実際：

A薬 100/2000 (5%)

B薬 200/2000 (10%)

<データベースにおけるアウトカム特定の  
感度が50%・特異度が80%の場合>

- 観測：

A薬 430/2000 (21.5%)

B薬 460/2000 (23%)

*Real-World??*

# Validation studyの種類

## 目的

- Target : 疾患・処置・薬剤・アウトカム
- 求める指標 : 感度／特異度  
陽性的中率／陰性的中率

## 方法

- カルテレビュー
- より信頼性が高いデータとの結合  
例：院内がん登録・症例データベース
- 他の情報からの“間接証拠”  
例：先行研究の発生率

# Validation studyの難しさ

van Walraven et al. Administrative database research infrequently used validated diagnostic or procedural codes. *J Clin Epidemiol* 2011 より抜粋：

- Measuring code accuracy can be time consuming and expensive.
- Securing grant support for such a study is difficult because a code-validation study may hold little appeal for many granting agencies.
- Finally, journals may decide to remove data regarding code validation in an effort to save space.

# 研究概要 (1)

## 概要：電子カルテをgold standardとした比較対象

- 国立病院機構の4病院
- 2014年度入退院、入院時18歳以上、DPC入院
- 各施設100入院をランダムに抽出

## 内容

- 病名： Charlson Comorbidity Indexの基準となる  
17疾患 (主or入院時併存として有無)
- 処置： 10種類の処置 (入院日の実施有無)
- 検査： 13種類の検査値 (入院日最初の値)

# 研究概要 (2)

## カルテレビュー方法

- 2名（看護師+医師）が独立に判断、相違があれば協議

## データベースからの抽出方法

- 病名： 様式1      主傷病, 入院の契機となった傷病,  
入院時併存病名×4      (疑い除く)
- 処置： EFファイル      入院日実施有無
- 検査： SS-MIX2ストレージ      入院日最初の値

## 解析

- 至適基準の質：  $\kappa$ 係数 (病名について)
- Validity： 感度・特異度、陽性・陰性的中率 (病名・処置)  
一致率 (検査値)



# 研究の実際 (2)

## カルテレビューの実施

- 日程
  - 連続する2日間で実施
  - 共同研究者・施設側との日程調整
- その他
  - 電子カルテの使用権限、環境の確保
  - 慣れない電子カルテの利用
  - 限られた時間
  - 再訪問は困難
  - レビュー結果データのやりとり
  - 
  -



# 研究の実際 (3)

## データベース上の作業

- 解析テーブルのイメージ

病院	患者	性別	主病名	併存症1	併存症2	併存症3	併存症4
A	1	M	J13	E119	I10	E785	
A	2	F	C220	B182			

DPC 様式1

尿検査	細菌培養	X線	CT
1	1	1	1
0	0	0	1

EFファイル

白血球数	Hb	CRP	Na
12500	14.5	9.6	141
	12.1		138

SS-MIX2

# 研究の実際 (3)

## データベース上の作業

- データ定義・オーダー例

診療報酬点数コード D018

入院日に実施あり→1、入院日に実施なし→0



尿検査	細菌培養	X線	CT
1	1	1	1
0	0	0	1

- 確認事項：

実施している患者の割合、病名との整合性、etc

# 研究の実際 (3)


## データベース上の作業

- データ定義・オーダー例

JLAC10コード            2A990000001992053  
入院日に実施した検査のうち最初の検査の結果

- 確認事項：

平均値、分布、  
単位、etc



白血球数	Hb	CRP	Na
12500	14.5	9.6	141
	12.1		138

# Validation studyの今後

## NHOにおける研究

- ① 国立病院機構の診療情報データベース高度  
利活用手法開発のためのバリデーション研究

NHO全体

疾患名に着目・より詳細な基準

- ② 診療情報データベースにおける入院後  
アウトカムの特定手法の開発

小規模科研費

術後アウトカムに着目

# NHOのデータを用いた臨床疫学研究

## NCDAの活用

- 従来のDPCデータ・レセプトデータに加えて、取得することができるもの：
  - 検査値
  - バイタル
  - 細菌検査結果
  - 薬剤感受性結果
  - 入院中の食事の有無と内容
  - など
- より臨床研究に近い詳細な研究を、大規模データベースで実施可能

# NHOのデータを用いた臨床疫学研究

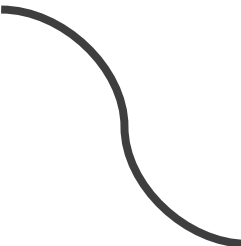
## 解析テーブルのイメージ

病院	患者	性別	主病名	併存症 1	併存症 2		手術1 code	手術1 day	
A	1	M	C187	E119	E785		K7193	4	
A	1	M	C187	E119	E785		K7193	4	
A	1	M	C187	E119	E785		K7193	4	
A	1	M	C187	E119	E785	...	K7193	4	...
A	1	M	C187	E119	E785		K7193	4	
A	1	M	C187	E119	E785		K7193	4	
A	1	M	C187	E119	E785		K7193	4	
A	1	M	C187	E119	E785		K7193	4	
A	2	F	C186	G470			K719-3	2	

# NHOのデータを用いた臨床疫学研究

## 解析テーブルのイメージ

1日1行のデータ



患者	手術1 day	day	体温	CRP	抗生剤	血培 MRSA	食事
1	4	1	36.9		0		1
1	4	2	36.4		0		1
1	4	3	36.5		0		1
1	4	4	38.1		1		0
1	4	5	38.5	15.1	1		0
1	4	6	38.0		1		0
1	4	7	37.8		1		0
1	4	8	37.6	19.8	0	0	0

# NHOのデータを用いた臨床疫学研究

## データ定義・オーダー例

- 体温

JLAC10コード： 9P1000000000000001 (体温計・腋窩),  
9P1500000000000001 (プローブ・直腸), etc

→ 1日の値のうちの最大値

- 抗生剤

HOTコード： 710618301, 710629901, etc

→ 1日にいずれか実施していれば1、なければ0

- 血液培養

JLAC10コード： 6B010000001974149

JANISコード： 1303 (MRSA)

→ 1日に1回でも検出されれば1、検出されなければ0